*Nucleic Acids Research, 2007, Vol. 35, Database issue* ***D1–D2***
*doi:10.1093/nar/gkl1051*

# EDITORIAL

The 2007 Database Issue of *Nucleic Acids Research* is the fourteenth in a series dedicated to databases in the field of molecular biology. These databases are essential resources for experimental and computational biologists alike and this compilation provides descriptions and updates of the most important of these databases, and serves to introduce newly compiled resources that provide specialist information in the biological area. The current issue is the largest yet and presents 68 new databases and updates of 106 existing databases. The 2007 Database Issue is not included in the print subscription to NAR. Instead, the Database Issue is freely available online to all under NAR's open access model. However, print copies are available for separate purchase by institutions and individuals.

Michael Galperin continues to produce and enlarge the Molecular Biology Database Collection, a compendium of databases that includes all those databases described in *Nucleic Acids Research*, as well as selected other databases relevant to biologists. NAR Online contains links to all of the databases in the compilation as well as brief summaries of their content. Individuals who wish to have their database listed in the Molecular Biology Database Collection or update a previous submission to the collection should contact Dr Michael Galperin directly (galperin@ncbi.nlm.nih.gov).

All authors wishing to submit articles for the 2008 Database Issue MUST contact Dr A. Bateman (nardatabase@mrc-lmb.cam.ac.uk) with a pre-submission enquiry, no later than July 1, 2007, to check whether a submission will be suitable for the issue. The pre-submission enquiry must present a working web accessible database for review by the Editor. Articles describing new databases will need to be received by August 15, 2007 at the latest, and should be prepared according to the instructions on the *Nucleic Acids Research* website (http://nar.oupjournals.org/). Authors who are submitting articles providing update information on databases that have previously been featured in *Nucleic Acids Research* should note that the deadline for submission of those articles is September 15, 2007.

The database issue would not be possible without timely reports from hundreds of reviewers. Thanks to you all! I would also like to thank Gill Smith for excellent editorial assistance. Finally I would like to thank Claire Saxby, and the rest of the team at Oxford University Press for producing this important issue.

## WHAT MAKES A GOOD DATABASE?

Having now edited the database issue for 4 years and carefully inspected over a thousand different biological databases. I feel I am well placed to give advice to prospective authors. There are many important aspects to any web accessible database that might be published in this issue. The quality, quantity and originality of data as well as the quality of the web interface are the most important. Good data with a poor interface or vice versa are never sufficient for consideration. Below I list a few pieces of advice that should make any future submission more likely to succeed as a resource and more likely to be published in the database issue.

### Data considerations

- When thinking of a name for your database do check if anyone else is using that name already. Calling your new database PDB is almost certainly going to cause confusion. It is also worth checking search engines with your database name, you may be surprised at what it means in other languages.
- Do make your data as comprehensive as possible. Try to avoid making the data collection overspecialized. For example, a database of promoters for RNA genes in a single organism is not going to have a wide appeal, but a database of promoters for RNA genes in all organisms would be of wide interest and utility.
- Do attribute the original sources of derived data.
- Do make sure that you are not breaching any license terms by redistributing data.
- Do include estimates of confidence in the data items if applicable.
- Do make data available for bulk download as flat files or relational database tables with associated documentation.
- Web services and DAS are becoming popular ways to make databases programmatically available. Making these available can stop your website being ground to a halt by users trying to screen scrape all your data.
- Do allow users to provide feedback on your data and submit new data. Do respond to user feedback in a timely manner.

### Web interface considerations

- Do document your web interface and data. Including a short tutorial will win you friends and influence reviewers.
- Do include a brief statement of what your database does on the front page.
- Make sure that users can always link back to the home page.
- Do not make all your links pop up in new windows.

- Do include example sequence/identifiers/keywords for every entry box on a query form.
- Keep search forms as simple as possible (think Google!) as most users will not want to do complex queries of your data. Keep advanced searches on a separate linked page.
- Do allow users to browse the data without searching for a specific entry. For example, provide alphabetical lists of entries, or entries sorted by function.
- Do not use a jargon term when a well-known term already exists.
- Make it obvious what information will be on a linked page and make clickable icons convey their function pictorially.
- Do get a domain name for your website. URLs to specific IP addresses/ports are unlikely to stand the test of time.
- Do test your website on a range of browsers and on a range of operating systems and make sure that external users can access all the content. Server errors, e.g. '404: page not found' errors are not popular with editors, reviewers or users.
- Do get feedback from your user community before submission, the more the better. Developers of databases are not always best placed to judge how easy their web interface is for users.
- Slow server response times will make your database unusable. You can help this by printing out at least some message to let users know that they should expect results in a certain time frame.
- There are many websites devoted to writing good web pages. Do look at these for pointers. Also have a look at the web pages describing web page faults to avoid such as http://www.webpagesthatsuck.com/.

If you are currently working on a new resource I hope that you will find these pointers useful. I look forward to seeing a pre-submission enquiry for your database soon.

Alex Batemen